Exploring the Evolution of the Gene Ontology and its Impact on Enrichment Analysis

Yi Chen[1,*], Frank W Takes[1], Fons J Verbeek[1] and Katherine J Wolstencroft[1]

[1]*Leiden Institute of Advanced Computer Science, Einsteinweg 55, 2333 CC, Leiden, The Netherlands*

Abstract

The Gene Ontology (GO) encapsulates shared knowledge on the functions of gene products and provides a knowledge structure for enrichment analyses of omics data. Such analyses enable researchers to place their findings in the context of current biological knowledge. These results are therefore time-sensitive, and change with our changing knowledge of biology. For research hot spots, such as cancer, or the recent SARS-CoV-2 pandemic, the evolution of knowledge is rapid and can heavily influence the interpretation and comparability of omics results. Consequently, as GO evolves, terms and their annotations are merged, added or made obsolete. This can have further influence on the meaning and consistency of conceptualized knowledge for related terms, which is known as semantic drift.

In this study, we investigated the extent of GO evolution and semantic drift by analysing changes to the GO network structure, connectivity and materialization. We assessed the impact of these temporal changes by reanalysing functional enrichment data and comparing the differences between biological conclusions that can be drawn with different versions of GO. In addition, we provide an open-source tool to enable researchers to perform functional enrichment analyses using GO from any time-point. It can be used to improve the comparability and reuse of published data, and to enable researchers to gain new insights from their own datasets as new knowledge is shared.

Keywords

Gene Ontology, Semantic Drift, Ontology Evolution, Functional Enrichment Analysis

1. Introduction

The Gene Ontology Knowledge Base (GOKB), made up of the Gene Ontology (GO) and its annotations (GOA) [1, 2] [3, 4] is used throughout the life sciences in many analysis methods, including functional enrichment, semantic similarity and link prediction[5, 6, 7, 8, 9, 10, 11]. These applications highlight the importance of the GOKB for analysing data in the context of our current biological knowledge. As our knowledge changes, GOKB evolves to reflect those changes, with some ontology terms and annotations becoming obsolete or merged, and new terms and annotations being added[12]. This means that reanalysing data with new versions of GOKB can provide new insights. Additionally, if studies are performed with outdated versions of GO, conclusions and results from these studies should not be directly compared[13].

Previous work has shown that researchers rarely provide metadata related to the version of GOKB in publications [14] and tools which use the GOKB as a knowledge source do not often

15th International Conference on Biological and Biomedical Ontology, July 17-19 2024, Enschede, The Netherlands

*Corresponding author.

y.chen@liacs.leidenuniv.nl (Y. Chen); f.w.takes@liacs.leidenuniv.nl (F. W. Takes); f.j.verbeek@liacs.leidenuniv.nl (F. J. Verbeek); k.j.wolstencroft@liacs.leidenuniv.nl (K. J. Wolstencroft)

follow the GO monthly release schedule, thereby providing analysis results based on outdated knowledge from older versions of GOKB[15]. This can have an impact on the reproducibility and comparability of results[16].

In this study, we conducted a comprehensive analysis of the changes to GOKB network structure from 2015 to 2021, aiming to understand the extent of ontology evolution (i.e. the explicit addition, merging or obsolescence of terms), and semantic drift (i.e. the implicit changes in the meaning of terms caused by the evolution of related terms). We measured the extent of these changes using changes to semantic similarity and ontology materialization and assessed the biological impact of GO evolution on differences in functional enrichment results. For functional enrichment, we present a case study from SARS-CoV-2 protein interaction data, where rapid comparison was necessary and time-sensitive, and different versions of GO lead to large differences in functional enrichment results. Overall, our research shows the importance of versioning for the use of GOKB, and motivates a greater focus on the temporal aspects of knowledge management. The tool we developed for performing the functional enrichment analysis and comparison is also provided at github, the tool dynamically collects GOKB data from any selected time-points, performs functional enrichment, and provides a method for visualizing the enriched terms according to their proximity in the GO hierarchy.

Materials and methods

Data Sources used in this study

- Gene Ontology and GO Annotation archive files were obtained from http://release.geneontology.org/ in monthly intervals from Jan 2015 to Dec 2021, and also from Sep 2023 (the latest version at the time of the analysis. We used human annotation files 'goa-human.gaf' which includes proteins annotated in Swiss-Prot[17] or the longest TrEMBL[18] transcript if there is no Swiss-Prot record.
- Protein reviewing status, describing UniProtKB[19] status of reviewed (Swiss-Prot) or unreviewed (TrEMBL), were acquired from UniprotKB[19]. A total of 20386 proteins were labelled reviewed, which denotes that they were annotated in Swiss-Prot.
- The SARS-CoV-2 viral protein interaction network was obtained from Gordon et al.[20] and can be downloaded from Ndex SARS-CoV-2. 332 high-confidence interactions between 27 SARS-CoV-2 viral proteins and human proteins were included.

GO and GO Annotations

A given version of the GO graph, G, can be represented as $G = (V, E)$ where V represents the set of GO terms and E represents edges in the graph. Edges are the relationships between two terms in the hierarchy, representing the majority of interactions in GO. GOA refers to all the statements relating GO terms to a particular gene product. For a single statement, a gene product is annotated with a specific GO term, supported by different sources represented by evidence codes, ranging from high quality manually annotated experimental evidence, to lower quality automatically generated annotations [21].

The number of terms and edges in GO were recorded for each release between Jan 2015 and Dec 2021. Differences between the numbers of new terms and obsolete terms were calculated using two consecutive releases of GO. We applied the same methodology to calculate the differences in GOA, and represented the differences by ratios. Ratios that changed by more than 10% across two consecutive releases were considered significant and further analysed to identify changes to Evidence Code (EC) composition[22]. The proportion of annotations categorized by each evidence code was also recorded. Increases in the number of manually curated evidence codes, as opposed to those automatically generated, were considered indicators of an increase in annotation quality.

Quantification of changes to the GO network

We analysed and characterized each GO aspect Biological process (BP), Molecular Function (MF), and Cellular Component (CC) separately, classifying edges into two groups based on their relationship type: subclass (is_a), and others (part_of, regulates etc). Then we computed the average degree for each aspect and the overlap ratio of the edge lists from Jan 2015 and Dec 2021, using a Jaccard index method (shown below).

$$J(V_1, V_2) = \frac{|V_1 \cap V_2|}{|V_1 \cup V_2|} \tag{1}$$

Where V_1, $V_2 \subset V$ and contain the top 20 nodes in 2015 and 2021 by degree.

We also calculated the degree centrality of GO terms to identify structurally important hub terms, as they play a critical role in network structure and knowledge representation. We compared the top 20 hub terms (i.e. those with the highest degree centrality scores) between Jan 2015 and Dec 2021, to explore the overall structural effects of GOKB evolution.

Impact of GO Evolution on Materialization

Materialization is the process of inferring implicit statements in an ontology based on provided axioms [23]. To evaluate the impact of GOKB evolution on the materialization of a specific term, we used a size-based metric to calculate the Consistency Score of the represented Biological Knowledge (CSBK).

$$CSBK = \frac{|A_1 \cap A_2|}{|A_1 \cup A_2|} \tag{2}$$

Where A_1 denotes the set of entailed axioms related to term v in 2015, and A_2 denotes the set of entailed axioms in 2021.

We use breadth-first search (BFS)[24] to find all possible paths to represent the axioms where the source was a particular term v in a given version of GO. If there was a path from a particular term v to an parent term v_p, then the subclass axioms linking the particular term v to v_p must exist. Terms were divided into three groups based on their positions; 1) internal terms, 2) leaf terms, and 3) Switch terms. Switch terms were defined as terms that transitioned from being internal in 2015 to leaf terms in 2021, or vice versa. For each group, we visualized the consistency scores by using a violin plot. The higher the consistency score, the less influence the terms received from the evolution of GOKB.

Assessing GO Evolution with Semantic Similarity

Semantic similarity methods measure the proximity of two ontology terms. Where semantic drift occurs, the proximity can change, resulting in larger or smaller values of semantic similarity. There are many different methods to calculate the semantic similarity between ontology terms. These include edge-based methods, which are based on the length of the path between compared terms[25], node-based methods, which rely on the Information Content values (e.g. Resnik[26][27]), and representation learning methods like Onto2Vec[28] and Node2Vec[29], which rely on the inferred and asserted logical axioms.

We used two methods to calculate the semantic similarity between GO terms in different GO versions. The first method, Resnik[26], is one of the most commonly used methods, and computes the semantic similarity of two ontology terms using the Information content (IC) value of their Most Informative Common Ancestor (MICA), that is:

$$sim(v_1, v_2) = IC(MICA) \tag{3}$$

The IC value is an indicator of the specificity and informativeness of a given term[26]. The IC value of a specific term can be determined by how frequently it appears within the annotation corpus, Which is:

$$IC(v) = -\log(p) \tag{4}$$

The IC value of the term can be normalized by dividing the maximum IC value. That is:

$$IC_{\text{norm}}(v) = IC(v)/IC_{max} \tag{5}$$

The second method we used was a machine learning-based method named Onto2Vec. Onto2Vec is an approach that learns the vector representation of GO by transforming the existing and inferred ontology axioms into text sequences and then embedding these ontology terms into vectors based on the context provided by the surrounding text[28]. Here, we used the Jan 2015 and Dec 2021 versions of GO and Onto2Vec to represent ontology terms by numerical vectors, using the python package mowl[30] version 0.3.0. The following parameters were used in Onto2Vec; Vector Size 200, Min Count 1, Window 10, Epochs 5, sg 1, negative 5.

To examine changes to semantic similarity, we randomly generated 20 sets of 50000 GO-term pairs, for each GO aspect that were not marked as obsolete in 2015 or 2021.

We used a paired t test to identify significant semantic similarity changes with the Benjamini-Hochberg multiple testing correction. Results with an adjusted p-value below 0.05 were classified as statistically significant. To ensure robustness, we repeated this experiment 5 times.

The Impact of GO Evolution on Functional Enrichment Analysis

We investigated the impact of GO evolution on enrichment analysis results using data from a highly cited SARS-CoV-2 Protein-Protein interaction network (PPI) published during the pandemic[20]. Since the outbreak of SARS-CoV-2, more than 960,000 research papers have been published on this topic, according to the Covid Data Portal (https://www. covid19dataportal.org/) [35]. This intense activity contributed to a rapid increase in biological knowledge. We therefore expected to observe changes in GOKB as a result, particularly in to immunological terms and

annotation. Our original intention was to compare pre- and post-pandemic GOKB versions and show the extent to which knowledge changed in that period. However, the original analysis showed that GO enrichment analysis was conducted using GOKB from MSigDB v6.1 (which was released in Oct 2017). Consequently, we analysed the data from that time-point in addition to using a GOKB version from a date close to the publication date (March 2020), and from a version after the pandemic (September 2023). We hypothesized that the use of an earlier version of GOKB in the original analysis may have resulted in missing insights relating to SARS-COv2 knowledge gained in the early days of the pandemic. Analysing three time-points allows us to investigate whether this was the case.

Enrichment analysis was conducted using a package we developed (V1.0), which can be found at here. The GOKB data used in the analysis was downloaded from the GO archive. We used the whole GO annotation as the reference set, and a hypergeometric test to calculate the p-values, and the adjusted p-values using Benjamini-Hochberg procedure (enriched terms <0.05). The top 20 enriched terms were used for comparing the differences between enriched results sets, using a Jaccard index and visualised using networkx version 3.2.1. The enriched terms of a viral protein cluster were grouped based on their proximity in the GO hierarchy. For any two enriched terms, if the length of the path between them was less than four, edges on the path were added to a networkx graph. Graph clusters were annotated with their Most Informative Common Ancestor (MICA).

Results

GO evolution - Terms and Annotations

Between Jan 2015 and December 2021, the number of terms in the Gene Ontology grew from 40,470 to 43,789. Fig 1A shows that instead of steady growth, the number of GO terms peaked in Dec 2018 and then started to decline until the end of 2021. Fig 1b presents the differences in the numbers of terms in two consecutive releases. We observed a greater variation in BP compared to the other two aspects, with more terms being added and being made obsolete.

A similar situation was observed for GO human annotations, which increased from 441,062 in Jan 2015 to 616,308 in Dec 2021. As Fig 1B and Table 1 shows, multiple reductions and increases occurred. The largest reduction occurred between June 2016 and July 2016, during which 89,743 annotations were removed, while only 260 annotations were added. This corresponds with the documented removal of annotations associated with unreviewed proteins. Analyzing the changes to the proportion of annotations categorized by different evidences codes in GOA (Fig 2) shows experimental evidence codes increased from 29.5% in Jan 2015 to 55% in Dec 2021, and became the largest annotation group, while electronic annotations, which constituted the largest part of GOA and accounted for 43.2% in 2015, experienced a decline from 190,464 in Jan 2015 to 73,882 in Dec 2021, ranked third by the end.

Quantifying changes to the GO network structure

Analysing the number edges in GO revealed an overall growth (Table 2) with overlaps in edge composition between 2015 and 2021 of 65% BP, 65% MF, and 35% CC, indicating large structural

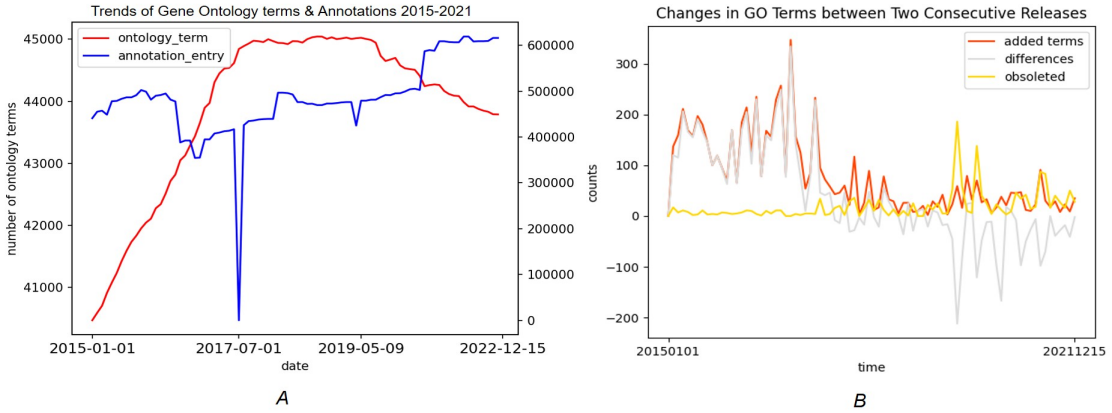

Figure 1: A) Number of GO terms and annotations per month, between Jan 2015 and Dec 2021. Due to the absence of annotation file for July 2017, there is a dip in annotations. B) Changes to GO terms between consecutive releases. Red = added terms, yellow = obsolete terms, and grey = total number of terms

changes across the network. The average degree, which is an indicator of the density of the network, decreased overall, with the largest being in CC (18%).

The degree distribution of GO follows a power law, and can therefore be considered scale-free. As such, the hub terms play an important role in connectivity. Significant changes to the highest ranked hub terms can signify large changes to the network structure. Our analysis showed the majority of CC hub terms changed (17/20), and one third of BP hub terms (7/20). There was a much smaller effect observed in MF hub terms (4/20).
Overall, we conclude that GOKB evolution exerts a large influence on the structure of GO.

Materialization of Biological Knowledge in GO

The conceptualized biological knowledge of a particular GO term can be seen as a specialisation or an intersection of biological knowledge represented by all its parents, according to asserted and inferred axioms. As parent terms become obsolete, or new parent terms are added, this conceptualized knowledge changes, influencing both asserted and inferred axioms[23], and

Table 1
GOA changes between 2015 and 2021. Annotation changes between 2015 and 2021. For two consecutive releases, a change ratio of 10% or more was considered significant

GO Version	Differences	Obsoleted	Added	Ratio
16.06.01–16.07.01	89,483	89,743	260	18.73%
16.11.01–16.12.01	39,877	44,867	4,990	11.23%
18.02.01–18.03.06	57,264	3,914	61,178	13.03%
19.03.18–19.04.17	51,381	3,989	55,370	10.79%
19.04.17–19.05.09	54,471	1,920	56,391	12.82%
20.06.01–20.07.16	88,470	4,500	88,970	16.80%

Table 2
GO Network Metrics

	Biological Process			Molecular Function			Cellular Component		
Parameter	2015	2021	pt[1]	2015	2021	pt	2015	2021	pt
terms	26,929	28,429	+5.57%	9,859	11,177	+13.4%	3,682	4,183	+13.61%
edges	65,515	66,348	+1.27%	12,428	14,003	+12.7%	7,182	6,997	+2.58%
avg_deg	4.87	4.67	-4.11%	2.52	2.51	-0.4%	3.90	3.20	-17.95%
avg_deg(is_a)	3.90	3.67	-5.90%	2.44	2.44	0%	3.07	2.30	-25.08%
avg_deg(other)[2]	0.97	1	+3.09%	0.08	0.07	-12.5%	0.83	0.9	+8.43%
overlap[3]	100%	65.09%	-34.91%	100%	65.49%	-34.51%	100%	34.74%	-65.26%

resulting in semantic drift.

Terms located at the leaf positions within the GO network have no child terms, so any modifications to them will solely influence the knowledge they themselves represent. Modifications to internal terms, however, especially those with a high degree, can greatly affect conceptualized knowledge (materialization) in the network. When analysing terms which became obsolete in our time-period, we found 49.77% of BP terms, 40.28% of MF terms and 45.16% of CC terms were internal, and 39.02% BP, 8.12% MF and 20.82% CC were added as new internal terms. This observation indicates large changes to the materialization of GO. To quantify the impact, we used a size-based metric from Pernisch et al. [23] to calculate the consistency score of the represented biological knowledge (CSBK) of all GO terms. Fig 3 shows a violin plot of these results. The median consistency scores were less than 0.4 in all three GO aspects. The average consistency score of BP, which occupied the largest proportion of terms, was around 0.8. It should be noted that Fig 3 also shows that not all terms were influenced by GOKB evolution in materialization. For MF, more than 50% of terms were stable, with a consistency score of 1.

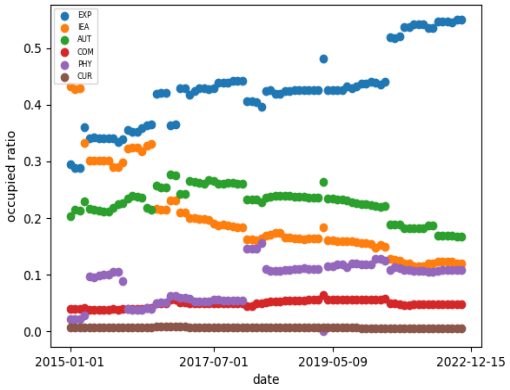

Figure 2: The ratio of annotation with different evidence codes. Experimental evidence codes (EXP), Phylogenetically-inferred annotations (PHY), Computational analysis evidence code (COM), Author statement evidence codes (AUT), Curator statement evidence code (CUR) and Electronic Annotation (IEA)

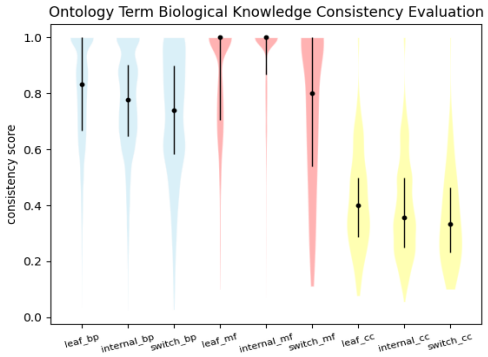

Figure 3: Knowledge Consistency Evaluation. Distribution of consistency scores in GO BP, MF and CC, classified into 3 groups: internal, in 2015 and 2021, leaf, in 2015 and 2021, and switch. Black dots are median values.

Assessing Changes to GO via Semantic Similarity

The proximity of terms in an ontology can me measured using semantic similarity. As an ontology evolves over time, we can therefore use semantic similarity to assess proximity changes. We randomly generated 20 sets of 50000 GO term pairs for each GO aspect and calculated the semantic similarity scores, using the Resnik and Onto2Vec methods, with Jan 2015 and Dec 2021 versions of GOKB, and then used a paired t-test to determine if there was a statistically significant difference. To ensure the robustness of the outcome, we repeated the experiment 5 times. For Resnik, all changes were significant. For Onto2Vec, table 3 shows all CC results and the majority of MF results were significant. For BP, over half were significant. The changes in proximity for the majority of terms indicate a large amount of semantic drift in GO over time.

Table 3
The paired t-test outcome of semantic similarity results, classified as significant (sig) <0.05 adj P-Value, or insignificant (Insig) >0.05 adj P-Value, repeated over 5 experiments.

	BP_Sig	BP_Insig	MF_Sig	MF_Insig	CC_Sig	CC_Insig
Exp1	14	6	20	0	20	0
Exp2	16	4	19	1	20	0
Exp3	13	7	17	3	20	0
Exp4	17	3	17	3	20	0
Exp5	16	4	19	1	20	0

The Influence of GO Evolution on Functional Enrichment Analysis

Functional enrichment analysis, over GOKB, is one of the most widely used methods for summarising and interpreting the outcome of large-scale differential expression experiments. As GO and its annotation corpus evolve, the terms that are considered enriched can change, which can change the interpretation of experimental conclusions. This can be a powerful way of

Table 4
Jaccard Index of top20 enriched GO terms for each SARS-CoV-2 protein and interacting proteins, comparing two periods: September 2017 to March 2020, and September 2017 to September 2023

Group	2017 vs 2020	2017 vs 2023	Group	2017 vs 2020	2017 vs 2023	Group	2017 vs 2020	2017 vs 2023
Nsp9	0.481	0.025	M	0.600	0.290	Orf8	0.538	0.481
Nsp2	0.905	0.111	Nsp1	0.818	0.379	Nsp12	0.379	0.538
Nsp13	0.739	0.176	Orf9b	0.538	0.379	Orf9c	0.538	0.600
Orf10	0.739	0.212	E	0.429	0.379	Nsp4	0.6	0.667
Nsp8	0.481	0.250	Nsp7	0.739	0.429	N	0.538	0.667
Orf3a	0.667	0.290	Nsp10	0.818	0.429			

gaining new insights from evolving biological knowledge, but it can also lead to reproducibility and comparability problems if versions of GOKB are not reported well. We reanalysed functional enrichment data from a highly cited SARS-CoV-2 research paper [20], where knowledge was rapidly evolving due to the recent pandemic. Gordon et al. investigated interactions between SARS-CoV-2 proteins and networks of interactions with human proteins. We reanalysed the data using the same versions of GO and GOA as used originally (September 2017). We show these results alongside results from performing the same analysis with versions of GO and GOA that were contemporary with the paper publication date (March 2020), and finally with results from a more current version, post-pandemic (September 2023).

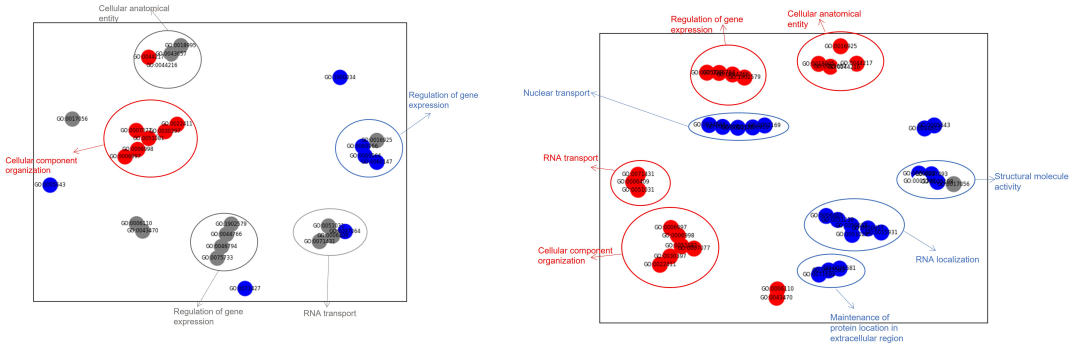

Figure 4: Top 20 enriched terms (ranked by adj P-Value) for SARS-CoV-2 viral bait protein Nsp9 and its interacting human proteins between 2017 and 2020 (left), and between 2017 and 2023 (right). Red and blue nodes exclusively enriched in earlier or later time-points respectively. Grey nodes enriched in both time-points. Clusters annotated with their most common ancestor.

Table 2 summarizes the results of a Jaccard index obtained by comparing the top 20 enriched terms between different time-points for each network cluster. Our re-analyses show that the majority of clusters have a score of below 0.5 when comparing enrichment between 2017 and 2023, which may be expected in a time-period of 6 years. However, four network clusters had scores of below 0.5 between 2017 and 2020. NSP9 is one example. Only 1 in 20 enriched terms were shared between 2017 and 2020. In Sep 2017, Nsp9 is enriched in terms related to RNA transport, cellular anatomical entities and cellular component organization, while at later

time-points, it was enriched in entities related to nuclear transport, structural molecule activity and RNA location, as figure 1 shows. If the original experiment had used an up-to date-version of GOKB for their analysis, the function of this cluster, and others with a low Jaccard score would have potentially been interpreted differently.

Discussion and Conclusion

The GOKB is a dynamic source of biological knowledge, which is essential for many omics data analyses. Our results highlight the importance of the temporal aspect of GO, and show the value of re-analysing datasets as our knowledge changes. Since 2015, the number of terms in GO has grown, but not continuously. There are periods of expansion, and periods of curation and consolidation. For GO annotation, there has been a general increase in experimentally confirmed annotations, and a reduction in automatically generated annotations. Both observations indicate an improvement to the quality of represented knowledge over time. An increase in the number of annotations may also indicate an increase in human protein function knowledge.

The network analysis results illustrate that GOKB evolution leads to changes in both the concepts and network structure, affecting the materialization process and inducing semantic drift. However, we focus specifically on human data, so it would be interesting to explore this phenomenon more broadly using data from other organisms.

The differences we observed between functional enrichment results were pronounced and would lead to differences in the biological interpretations of results. It was expected that GOKB would change rapidly in response to new SARS-Cov2 knowledge. However, the results were more striking than expected because the authors of the original paper used an outdated version of GOKB for their initial analysis. The reason for this choice was unclear. The authors reported the version of the analysis tool that was used, but not the version of GOKB. It is possible that it was used with the assumption it was the latest version of GOKB. In a related study[14], we found that some enrichment analysis tools did not report their GOKB update schedule or the version of GOKB they currently used. Therefore, it may not always be possible for researchers to easily determine whether analysis tools are using the latest knowledge. In situations where it is essential that the latest knowledge is incorporated, using the latest available versions of knowledge resources, and accurately reporting their versions is important for reproducibility.

The continuous evolution of GOKB reflects our evolving biological knowledge, so this dynamic component is essential for sharing continuous scientific advances and a collective understanding of biological processes. With increased awareness of the network effects of GOKB evolution and semantic drift, we can improve reproducibility and therefore comparability and reusability of such data across the life sciences.

References

[1] M. Ashburner, C. A. Ball, J. A. Blake, D. Botstein, H. Butler, J. M. Cherry, A. P. Davis, K. Dolinski, S. S. Dwight, J. T. Eppig, et al., Gene ontology: tool for the unification of biology, Nature genetics 25 (2000) 25–29.

[2] The gene ontology resource: enriching a gold mine, Nucleic acids research 49 (2021) D325–D334.
[3] G. O. Consortium, Gene ontology annotations and resources, Nucleic acids research 41 (2012) D530–D535.
[4] D. Barrell, E. Dimmer, R. P. Huntley, D. Binns, C. O'Donovan, R. Apweiler, The goa database in 2009—an integrated gene ontology annotation resource, Nucleic acids research 37 (2009) D396–D403.
[5] T. R. Dalmer, R. D. Clugston, Gene ontology enrichment analysis of congenital diaphragmatic hernia-associated genes, Pediatric research 85 (2019) 13–19.
[6] M. Yousef, A. Kumar, B. Bakir-Gungor, Application of biological domain knowledge based feature selection on gene expression data, Entropy 23 (2020) 2.
[7] P. Denny, M. Feuermann, D. P. Hill, R. C. Lovering, H. Plun-Favreau, P. Roncaglia, Exploring autophagy with gene ontology, Autophagy 14 (2018) 419–436.
[8] Y.-H. Zhang, T. Zeng, L. Chen, T. Huang, Y.-D. Cai, Determining protein–protein functional associations by functional rules based on gene ontology and kegg pathway, Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics 1869 (2021) 140621.
[9] I. Tripodi, K. B. Cohen, L. E. Hunter, A semantic knowledge-base approach to drug-drug interaction discovery, in: 2017 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), IEEE, 2017, pp. 1123–1126.
[10] M. Asif, H. F. Martiniano, A. M. Vicente, F. M. Couto, Identifying disease genes using machine learning and gene functional similarities, assessed through gene ontology, PloS one 13 (2018) e0208626.
[11] S. Althubaiti, A. Karwath, A. Dallol, A. Noor, S. S. Alkhayyat, R. Alwassia, K. Mineta, T. Gojobori, A. D. Beggs, P. N. Schofield, et al., Ontology-based prediction of cancer driver genes, Scientific reports 9 (2019) 1–9.
[12] R. P. Huntley, T. Sawford, M. J. Martin, C. O'Donovan, Understanding how and why the gene ontology and its annotations evolve: the go within uniprot, GigaScience 3 (2014) 2047–217X.
[13] A. Tomczak, J. M. Mortensen, R. Winnenburg, C. Liu, D. T. Alessi, V. Swamy, F. Vallania, S. Lofgren, W. Haynes, N. H. Shah, et al., Interpretation of biological experiments changes with evolution of the gene ontology and its annotations, Scientific reports 8 (2018) 1–10.
[14] Y. Chen, F. Verbeek, K. Wolstencroft, FAIR Functional Enrichment: Assessing and Modelling Provenance in Omics Results, in: Proceedings of the 14th International Conference, SWAT4HCLS, Basel, Switzerland, 2023, p. Forthcoming.
[15] L. Wadi, M. Meyer, J. Weiser, L. D. Stein, J. Reimand, Impact of outdated gene annotations on pathway enrichment analysis, Nature methods 13 (2016) 705–706.
[16] M. Jacobson, A. E. Sedeño-Cortés, P. Pavlidis, Monitoring changes in the gene ontology and their impact on genomic data analysis, GigaScience 7 (2018) giy103.
[17] E. Boutet, D. Lieberherr, M. Tognolli, M. Schneider, P. Bansal, A. J. Bridge, S. Poux, L. Bougueleret, I. Xenarios, Uniprotkb/swiss-prot, the manually annotated section of the uniprot knowledgebase: how to use the entry view, in: Plant Bioinformatics, Springer, 2016, pp. 23–54.
[18] B. Boeckmann, A. Bairoch, R. Apweiler, M.-C. Blatter, A. Estreicher, E. Gasteiger, M. J. Martin, K. Michoud, C. O'Donovan, I. Phan, et al., The swiss-prot protein knowledgebase

and its supplement trembl in 2003, Nucleic acids research 31 (2003) 365–370.
[19] Uniprot: the universal protein knowledgebase in 2021, Nucleic acids research 49 (2021) D480–D489.
[20] D. E. Gordon, G. M. Jang, M. Bouhaddou, J. Xu, K. Obernier, K. M. White, M. J. O'Meara, V. V. Rezelj, J. Z. Guo, D. L. Swaney, et al., A sars-cov-2 protein interaction map reveals targets for drug repurposing, Nature 583 (2020) 459–468.
[21] N. Škunca, A. Altenhoff, C. Dessimoz, Quality of computationally inferred gene ontology annotations, PLoS computational biology 8 (2012) e1002533.
[22] S. Nadendla, R. Jackson, J. Munro, F. Quaglia, B. Mészáros, D. Olley, E. T. Hobbs, S. M. Goralski, M. Chibucos, C. J. Mungall, et al., Eco: the evidence and conclusion ontology, an update for 2022, Nucleic acids research 50 (2022) D1515–D1521.
[23] R. Pernisch, D. Dell'Aglio, A. Bernstein, Beware of the hierarchy—an analysis of ontology evolution and the materialisation impact for biomedical ontologies, Journal of Web Semantics 70 (2021) 100658.
[24] C. Y. Lee, An algorithm for path connections and its applications, IRE transactions on electronic computers (1961) 346–365.
[25] V. Pekar, S. Staab, Taxonomy learning-factoring the structure of a taxonomy into a semantic classification decision, in: COLING 2002: The 19th International Conference on Computational Linguistics, 2002.
[26] P. Resnik, Semantic similarity in a taxonomy: An information-based measure and its application to problems of ambiguity in natural language, Journal of artificial intelligence research 11 (1999) 95–130.
[27] C. Zhao, Z. Wang, Gogo: An improved algorithm to measure the semantic similarity between gene ontology terms, Scientific reports 8 (2018) 15107.
[28] F. Z. Smaili, X. Gao, R. Hoehndorf, Onto2vec: joint vector-based representation of biological entities and their ontology-based annotations, Bioinformatics 34 (2018) i52–i60.
[29] A. Grover, J. Leskovec, node2vec: Scalable feature learning for networks, in: Proceedings of the 22nd ACM SIGKDD international conference on Knowledge discovery and data mining, 2016, pp. 855–864.
[30] F. Zhapa-Camacho, M. Kulmanov, R. Hoehndorf, mowl: Python library for machine learning with biomedical ontologies, Bioinformatics 39 (2023) btac811.

A. Online Resources

The source code for the functional enrichment tool can be found at github.com/chestnzu/GO-evolution.git.
Supplementary data can be found at fairdomhub.org/investigations/665